

DATA ANALYSIS IN COMMUNICATION

Spring 2010
3 credits
Tue. 2-5pm 大勇402

Instructor: Tsung-Jen Shih
tjshih@nccu.edu.tw

Course outline:

This is an intermediate course in quantitative data analysis for communication-related topics. The major objectives are to (a) help you understand the prerequisites for collecting data that are suitable for subsequent analysis, and to (b) analyze existing data sets using the appropriate tools.

Since this is an intermediate course in data analysis, some familiarity with key methods of data gathering (survey, content analysis, experiment, etc.) and statistics (univariate and multivariate) is expected. We will talk about many of the methods and statistical tools you will encounter, but not at an introductory level. Thus, previous coursework in statistics and methods is strongly advised.

The course is designed as a lab course. I will provide context and some lecture material, but the majority of course time will be spent analyzing data using various software packages, trying to make sense of output, and linking data to the research questions that they were supposed to test.

Course requirements:

All assignments are based on the assumption that one or two semesters down the road you will need what we have learned for a paper in another class, a conference presentation, a publication, or a thesis. Therefore, please prepare handouts of everything you do for all students in the class. This way, we will all have short summaries of the most important methods, their pitfalls, and of how to run them in SPSS or other statistical software packages.

(1) *Class participation* (10%):

I will grade your participation in class, including your work as a class discussant, team presenter, and paper discussant during the final paper presentations. For your contributions to seminar discussion. It is not enough that you just come to class. You are expected to actively discuss the readings or methodological problems and critically analyze their contents. In addition, during the last two weeks, each seminar participant is expected to present his or her research paper to the course, including a longer discussion of the methodological and statistical challenges you encountered in your study. Each paper will also be discussed by two other participants, similar to a conference presentation. For the presenters, this means that they should share their papers with their discussants at least **72 hours** before their presentation. The discussants, in turn, are expected to provide informed and critical feedback. Like all academic discourse, this feedback should be based on evidence and information rather than normative views and opinions.

(2) *Topic presentations* (20%):

Each student will sign up as a discussant for one or two weeks. This presentation should not focus on the statistical or mathematical background of the method but rather on its application. How can this method be used to answer a research question in communication science? Ideally, this research question should come directly out of your Master's project or other research that you are working on. It is recommended that you work with a data set that you will use for your seminar paper or for another project you are involved in.

For the methods presentation, please prepare

- a 1-2 page handout with a quick overview of the analytic tool you are presenting,
- two readings (one being the article or chapter you found most helpful as an overview, and one being a good application of this tool in the communication literature)
- the SPSS syntax of the analyses you showed in class.

Please send the materials for all your presentations **at least 48 hours before you present**.

Please also send an email to everyone once the materials are online.

(3) Conceptualization and operationalization (15%)

This assignment will help you understand how concepts are defined and measured in academic studies. Students will need to pick a concept (e.g. political participation, media use, science literacy) and go to communication journals (such as CR, JOC, PUS, Science Comm, POQ, Mass comm & society, IJPOR, etc.) for related studies (3-5). Then you should record information about

- 1). the relationship between variables
- 2). how concepts are defined
- 3). how concepts are measured (operationalization, reliability)
- 4). how numbers/ statistics are presented
- 5). the methodological advantage and drawbacks of your selected studies

(4) Overview of journals:

This is a very short assignment. Every student will examine two to three peer-reviewed academic or professional journals that could serve as outlets for his or her seminar paper. You will then present – in class – why you think these journals provide a good fit for what you're doing. Who is the editor, what is the acceptance rate, what has been published there in previous issues, how long can articles be, how are empirical analyses typically presented, etc.?

For each journal, provide a one-page handout for everyone in class. This assignment will not be formally graded, but of course it will go into the qualitative assessment of your class participation grade.

(5) Individual research project/presentation (55%):

Finally, what you have learned in the previous assignments will culminate in a research paper based on quantitative data analysis. Specifically, this will include outlining a problem and its relevance for a specific discipline, translating the problem into research questions and testable hypotheses, developing measures, and providing an analytic answer. In contrast to other seminar papers, feel free to provide appendices or additional materials to justify your analytic choices or show competing analytic approaches. The final paper is due on **June 22**.

All written assignments in this class should be formatted using 12-point font (Arial, Helvetica, or Times New Roman) and double line spacing, and follow a commonly accepted academic citation style (APA, Chicago, etc.). Please make also sure that all of your assignments live up to minimal professional standards, i.e., are stapled, have cover pages, page numbers, etc.

Readings:

Here is a list of good overview books and manuals. These should be your first stop when you are beginning to tackle a new analytic method or are looking for an introductory reading for a particular week.

Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Wadsworth.

Bryman, A. & Cramer, D. (2002). *Quantitative data analysis with SPSS 10 for Windows: A guide for social scientists*. New York: Routledge.
 Hayes, A. F. (2005). *Statistical methods for communication science*. Mahwah, NJ: Erlbaum.
 Kerlinger, F. N. (1986). *Foundations of behavioral research* (3rd edition). New York: Holt-Rinehart.
 Wimmer, R. D., & Dominick, J. R. (2003). *Mass media research: An introduction* (7th edition). Belmont, CA: Wadsworth.
 Levin J., & Fox J. A. (1988). *Elementary statistics in social research*. New York: Harper & Row.

Relevant Online Resources

SPSS manuals @ Syracuse:

<http://qsic.syr.edu/manuals/spss/>

Statistical computing @ UCLA (annotated output and manuals for various statistical software)

<http://www.ats.ucla.edu/stat/>

National Election Study data & questionnaire archive @ Michigan:

<http://www.electionstudies.org/>

Election Study Center @ NCCU

<http://esc.nccu.edu.tw/newchinese/data/database.htm>

The Pew Research Center for the People and the Press

<http://people-press.org/dataarchive/>

Eurobarometer

http://ec.europa.eu/public_opinion/index_en.htm

ICPSR

<http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp>

Schedule:

Week 1: 2/23 Introduction and overview of basic statistics

Data resources

- Where to download datasets
- How to deal with downloaded data

Week 2: 3/2 Introduction: Scientific Approaches and Social Science (Why do we do research the way we do?)

- Chalmers, A. F. (1982). *What is this thing called science?* Indianapolis, IN: Hackett. (Ideally you should read the whole book, but definitely Chapters 1-6.)
- McLeod, J. M. and Reeves, B. (1980). On the nature of mass media effects. In Stephen B. Withey and Ronald P. Abeles (Eds.), *Television and social behavior* (pp. 17-54). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tichenor, P. J., & McLeod, D. M. (1989). The logic of social and behavioral science. In G. H. S. III & B. H. Westley (Eds.), *Research methods in mass communication* (2nd ed., pp. 10-29). Englewood Cliffs, NJ: Prentice Hall.
- Popper, K. R. (1959). *The logic of scientific discovery*. London: Hutchinson.
- Reynolds, P. D. (1971). *A primer in theory construction*. New York: Macmillan.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd enlarged ed.).

- Chicago, IL: University of Chicago Press.
- Blalock, H. M. (1969). *Theory construction: From verbal to mathematical formulations*. Englewood Cliffs, NJ: Prentice-Hall.

Week 3: 3/9 From explication to measurement, and back: The logic of empirical social science research

Readings:

- Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Wadsworth. (Chapters 4-6)
- Eveland Jr., W. P. (1997). Interactions and nonlinearity in mass communication: Connecting theory and methodology. *Journalism & Mass Communication Quarterly*, 74(2), 400-416.
- McLeod et al. (1999). *Concept explication and theory construction, Parts I + II*. Working paper, University of Wisconsin-Madison.
- You may want to check out this simulation of sample, sampling, and population distributions, just as a refresher on distribution-based statistics:
http://onlinestatbook.com/stat_sim/sampling_dist/index.html
- Blalock, H. M. (1969). *Theory construction: From verbal to mathematical formulations*. Englewood Cliffs, NJ: Prentice-Hall. (Chapter 3, pp. 27-47).
- Chaffee, S. H. (1991). *Explication*. Newbury Park, CA: Sage.

Week 4: 3/16 Data cleaning and preparation, descriptive, missing values, etc.

Discussion topics:

- Basics of SPSS (How does the software work?)
- Syntax, output, and data windows
- Changing SPSS Preferences
- Syntax language and pull-down menus
- Reading in raw data, and SAV and POR files
- Value and variable labels
- System vs. user-defined missing values
- etc.
- Getting a first overview of the data set (demographics, media use measures, measures of political participation, etc.)
- Descriptives (mean, median, standard deviations, non-normality)
- Frequency tables (Valid and total percentages)
- Simple recoding (COMPUTE, RECODE and IF commands)

Week 5: 3/23 Measurement decisions and evaluation

- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Newbury Park, CA: Sage.
- Rubin, R. B., Palmgreen, P., & Sypher, H. E. (1994). *Communication research methods: A source book*. New York: Guildford.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). *Measures of personality and social psychological attitudes*. New York: Academic Press.

Discussion topics:

- How to create weighted vs. unweighted scales –
- Summative/averaged scales in SPSS (problems with missing values, scale ranges, etc.)
- Reliability testing, validity testing in SPSS (look over the Concept Explication chapters from last week for this)

Week 6: 3/30

Assignment due: conceptualization and operationalization

Exploratory factor analysis

Discussion topics:

- Weighted vs. unweighted scales
- Extraction and rotation methods
- Forced solutions vs. extraction based on Eigenvalues

- Week 7: 4/6 Analysis of (quasi) experimental data
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin.
 - Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Boston, MA: Houghton Mifflin.
 - Iversen, G. R., & Norpoth, H. (1987). *Analysis of Variance*. Newbury Park, CA: Sage.
 - Scheffé, H. (1959). *The analysis of variance*. New York: Wiley.
- Week 8: 4/13 Introduction to Regression
- Achen, C. H. (1982). *Interpreting and using regression*. Newbury Park, CA: Sage.
 - Lewis-Beck, M. S. (1980). *Applied regression: An introduction*. Newbury Park, CA.
 - Schroeder, L. D., Sjoquist, D. L., & Stephan, P. E. (1986). *Understanding regression analysis: An introductory guide*. Newbury Park, CA: Sage.
- Discussion topics:**
- The logic of regression
 - OLS vs. Logistic
 - Regression assumptions
 - R-squares, betas, and Bs
 - How to report regressions
- Week 9: 4/20 Hierarchical Regression/ test of mediation
- Discussion topics:**
- Different types of betas
 - R square change
 - Arrangement of different variables
 - Spurious relationships
- Final paper prospectus due**
- Week 10: 4/27 Interactions in regression
- Discussion topics:**
- Creating interaction terms
 - Main effects and interaction effects
 - Plotting interactions
 - Interactions and nonlinearity
- Week 11: 5/4 Regression recap/ working session
- Week 12: 5/11 Regression path analysis
- Week 13: 5/18 Hierarchical linear modeling
- Week 14: 5/25 Develop a scholarly paper
- APA style
 - Paper structure
 - Etc...
- Week 15: 6/1 Preparation of manuscripts and publishing
- Readings:**
- ISI Web of Knowledge - Journal Citation Report:
<http://ezproxy.library.wisc.edu/login?url=http://digital.library.wisc.edu/1711.web/wok>
 - Eigenfactor.org:
<http://www.eigenfactor.com/map/>

- Campbell, P. (2008). Escape from the impact factor. *Ethics in Science and Environmental Politics*, 8, 5-7.

Discussion topics:

- Preparing articles for publication
- Overview of relevant journals and conferences
- Evaluating the quality of journals and their relevance for your career path
- The review process

The ethics of data analysis, presentation, and scholarly publishing

Assignment due: journal review

Week 16: 6/8 Final paper presentation

Week 17: 6/15 Final paper presentation